

MEASURING DARK MATTER PROFILES NON-PARAMETRICALLY IN DWARF SPHEROIDALS: AN APPLICATION TO DRACO

JOHN R. JARDEL¹, KARL GEBHARDT¹, MAXIMILIAN FABRICIUS², NIV DRORY³, AND MICHAEL J. WILLIAMS²

ACCEPTED TO APJ: 11 November 2012

ABSTRACT

We introduce a novel implementation of orbit-based (or Schwarzschild) modeling that allows dark matter density profiles to be calculated non-parametrically in nearby galaxies. Our models require no assumptions to be made about velocity anisotropy or the dark matter profile. The technique can be applied to any dispersion-supported stellar system, and we demonstrate its use by studying the Local Group dwarf spheroidal (dSph) galaxy Draco. We use existing kinematic data at larger radii and also present 12 new radial velocities within the central 13 pc obtained with the VIRUS-W integral field spectrograph on the 2.7m telescope at McDonald Observatory. Our non-parametric Schwarzschild models find strong evidence that the dark matter profile in Draco is cuspy for $20 \leq r \leq 700$ pc. The profile for $r \geq 20$ pc is well-fit by a power law with slope $\alpha = -1.0 \pm 0.2$, consistent with predictions from Cold Dark Matter (CDM) simulations. Our models confirm that, despite its low baryon content relative to other dSphs, Draco lives in a massive halo.

Subject headings: dark matter—galaxies: dwarf—galaxies: individual (Draco)—galaxies: kinematics and dynamics—Local Group

1. INTRODUCTION

Understanding how dark matter is distributed in low-mass galaxies is central to the study of galaxy formation in the cold dark matter (CDM) paradigm. The first CDM simulations predicted that all dark matter halos share a universal density profile with a cuspy inner slope of $\alpha \equiv d \ln \rho_{DM} / d \ln r = -1$ (Navarro et al. 1996b, hereafter NFW). When observers began studying low-mass galaxies, however, they mostly found halos with a uniform density $\alpha = 0$ core (Burkert 1995; Persic et al. 1996; Borriello & Salucci 2001; de Blok et al. 2001; Blais-Ouellette et al. 2001; Simon et al. 2005). This disagreement between theorists and observers over the form of $\rho_{DM}(r)$ became known as the core/cusp debate.

Since the debate began, the number of profile parameterizations used to describe low-mass galaxies by both theorists and observers has only increased. Observers champion empirical fits such as the Burkert profile (Burkert 1995; Salucci & Burkert 2000), cored isothermal models (Persic et al. 1996) or simply generic broken power laws (Koch et al. 2007; Strigari et al. 2008; Walker et al. 2009). Theorists have also introduced new fits to their simulated halos with varying, although still cuspy, inner slopes (Navarro et al. 2004; Stadel et al. 2009; Navarro et al. 2010). Modeling a galaxy with each of these parameterizations would not only be time consuming, but also biased if the true profile is unlike any of them. It is therefore preferable to use non-parametric methods to determine $\rho_{DM}(r)$.

Van den Bosch et al. (2006) first experimented with non-parametric orbit-based models by allowing the mass-to-light ratio M/L to vary with radius in the globular cluster M15. We introduce a similar modeling technique that uses axisym-

metric Schwarzschild modeling, combined with knowledge of the full line-of-sight velocity distribution (LOSVD) of stars, to break the well-known degeneracy between mass and orbital anisotropy. We demonstrate the capability of these models by applying them to the Local Group dwarf spheroidal (dSph) galaxy Draco. Draco is part of an interesting class of objects that are some of the most dark matter-dominated galaxies discovered. This makes differentiating between dark and luminous mass in dSphs easier as the baryons have less of an effect on the total density profile than they do in larger galaxies. Recently, using improved data and modeling techniques, Adams et al. (2012) found a cuspy dark matter profile in the low-mass galaxy NGC 2796 where previous studies found a core. Studies like these motivate us to investigate the dSphs with more sophisticated models.

Our models represent a significant improvement over previous work on dSphs as most studies use spherical Jeans models (Gilmore et al. 2007; Walker et al. 2009; Wolf et al. 2010) which require the modeler to make assumptions about the nature and degree of the anisotropy. These assumptions vary in complexity from simply assuming isotropy, which can bias the estimate of α (Evans et al. 2009), to parameterizing the anisotropy as a general function of radius (Strigari et al. 2008; Wolf et al. 2010). Models that allow for more freedom in the anisotropy typically fall victim to the mass-anisotropy degeneracy and cannot robustly determine the inner slope of $\rho_{DM}(r)$ (Walker et al. 2009). We hope to make a robust determination of the inner slope in Draco with a suite of more general non-parametric Schwarzschild models.

2. NON-PARAMETRIC SCHWARZSCHILD MODELS

At the heart of our non-parametric technique is the orbit-based modeling code developed by Gebhardt et al. (2000, 2003), updated by Thomas et al. (2004, 2005), and described in detail in Siopis et al. (2009). All orbit-based codes are based on the principle of orbit superposition first introduced by Schwarzschild (1979). Similar axisymmetric codes are used by Rix et al. (1997), van der Marel et al. (1998), Cretton et al. (1999), and Valluri et al. (2004) while van den Bosch et al. (2008) present a fully triaxial modeling

¹ The University of Texas, Department of Astronomy, 2515 Speedway, Stop C1400, Austin, Texas 78712-1205; jardel@astro.as.utexas.edu

² Max-Planck Institut für extraterrestrische Physik, Giessenbachstrasse, D-85741 Garching bei München, Germany

³ Instituto de Astronomía, Universidad Nacional Autónoma de México, Avenida Universidad 3000, Ciudad Universitaria, C.P. 04510 México D.F., México

code. The current Schwarzschild models that allow for dark matter do so by requiring the modeler to assume a parameterization for the dark matter density profile $\rho_{DM}(r)$. Unfortunately, this parameterization is often exactly what we wish to determine. Current methods get around the circular logic of this dilemma by running models with different parameterizations and comparing their relative goodness-of-fit with a χ^2 test. Non-parametric modeling sidesteps the issue entirely, and lets the parameterization of $\rho_{DM}(r)$ be output from the models, rather than input as a guess.

The principle of orbit superposition works by choosing from a library of all possible stellar orbits only those orbits that best reproduce the observed kinematics of the galaxy being modeled. If we know the mass density profile of the galaxy, and hence the potential, we can compute the appropriate orbit library. However, since we do not know the potential of the galaxy, we must construct a number of models with slightly different mass distributions and compare the goodness-of-fit of the resulting allowed orbits. The radial profile of the total (dark + stellar) mass density in a galaxy can be written as:

$$\rho(r) = \frac{M_*}{L} \times \nu(r) + \rho_{DM}(r) \quad (1)$$

where M_*/L is the mass-to-light ratio of the stars, $\nu(r)$ is the stellar luminosity density profile, and $\rho_{DM}(r)$ is the dark matter density profile. In principle we know M_*/L , which can vary as a function of radius, from stellar population models. We also know $\nu(r)$ from the de-projection of the observed surface brightness profile. Our task is to construct orbit libraries for varying $\rho(r)$ and match the allowed orbits to kinematics in the form of LOSVDs—the distribution of projected velocities observed. Some orbit libraries will contain orbits that do a good job at fitting the observed LOSVDs and others will not. The best-fitting model identifies the best-fitting $\rho(r)$. Once we know this, we can invert Equation (1) to solve for $\rho_{DM}(r)$. The trick is to vary $\rho(r)$ in a systematic way. This is the principal difference between our new approach and standard Schwarzschild modeling which tries to vary $\rho(r)$ by varying the parameters that define an assumed dark matter profile.

To compute the orbit library for each model, we first calculate the potential. We assume axisymmetry and make use of the stellar ellipticity to define the density at angle θ in our meridional grid. The dark matter halo is assumed to have the same ellipticity as the stars. We solve Poisson's equation for the potential associated with this density distribution by decomposing $\rho(r, \theta)$ into spherical harmonics (Siopis et al. 2009). With the potential known, we launch 20,000-30,000 orbits and integrate their motion for roughly 100 crossing times, storing position and velocity information at each timestep.

Orbits in axisymmetric potentials respect three isolating integrals of motion: energy E , the z -component of angular momentum L_z , and the non-classical third integral I_3 . By specifying all three of these quantities together, an orbit is uniquely defined. Unfortunately, there is no analytical form for I_3 and it is generally not known a priori. We therefore rely on the sampling scheme of Thomas et al. (2004) to construct an orbit library which uniformly samples E , L_z , and I_3 and thereby contains all possible orbits for a given potential.

Each orbit in the library is given a weight w_i , and a set of w_i are chosen so the observed kinematics are appropriately reproduced by the orbits which have been weighted, averaged,

and projected. Quantitatively, we observe N_{LOSVD} LOSVDs in the galaxy at various positions. Each LOSVD contains N_{vel} velocity bins with uncertainties, so the number of observables the models must match to is given by the product $N_{\text{LOSVD}} \times N_{\text{vel}}$. The goodness-of-fit of a model is judged by

$$\chi^2 = \sum_{i=1}^{N_{\text{LOSVD}}} \sum_{j=1}^{N_{\text{vel}}} \left(\frac{\ell_{ij}^{\text{obs}} - \ell_{ij}^{\text{mod}}}{\sigma_{ij}} \right)^2 \quad (2)$$

where ℓ_{ij}^{obs} and ℓ_{ij}^{mod} are the j^{th} velocity bin of the i^{th} LOSVD from the observations and model respectively, and σ_{ij} is the uncertainty in ℓ_{ij}^{obs} .

Given the freedom to choose from upwards of 10,000 orbital weights to match only $N_{\text{LOSVD}} \times N_{\text{vel}} \sim 100$ observables, a standard χ^2 minimization routine can populate the distribution function in any number of ways that introduce unwanted noise. To avoid distribution functions that are noisy or unrealistic, however still consistent with the observables, we employ a maximum entropy smoothing technique developed by Richstone & Tremaine (1988) and described in Siopis et al. (2009). Instead of minimizing χ^2 , we maximize the objective function

$$\hat{S} = - \sum_{i=1}^{N_{\text{orb}}} w_i \log \left(\frac{w_i}{\Delta\Omega_i} \right) - \alpha_S \chi^2 \quad (3)$$

where N_{orb} is the number of orbits in the library, and $\Delta\Omega_i$ is the phase-space volume of the i^{th} orbit. See Siopis et al. (2009) for a technical description of how we calculate phase-space volumes and maximize \hat{S} .

The first term in Equation (3) is an entropy-like quantity and the second term is χ^2 from Equation (2). The parameter α_S controls which term influences \hat{S} . For small α_S , orbital weights are chosen to produce a smooth distribution function at the expense of reproducing the data. For large α_S , the data are well-fit by the model (χ^2 is small), but the resulting distribution function is likely not smooth. We determine the appropriate α_S for each model using the scheme described in Siopis et al. (2009). We start with $\alpha_S = 0$ and incrementally increase it until changes to χ^2 between successive iterations are small. Thus, the maximum entropy constraint serves to initialize the search for the minimum when $\alpha_S = 0$. By slowly increasing α_S , we drive down the importance of entropy to the fit until it no longer matters.

2.1. Varying $\rho(r)$ Between Models

The major innovation of our new modeling technique is how we choose the density profile $\rho(r)$ of each model. Current methods assume $\rho_{DM}(r)$ and calculate $\rho(r)$ from Equation (1), however this requires knowledge of the appropriate parameterization for $\rho_{DM}(r)$. We use a fundamentally different strategy and divide $\rho(r)$ into N_{bin} discrete points whose value ρ at radius r_i is labeled ρ_i . The N_{bin} points are spaced evenly in $\log r$ and connected by straight line segments. Each trial density profile is now defined by the ρ_i at each of the N_{bin} bins. We run many models adjusting the values of the ρ_i so as to sample all possible density profiles. This strategy requires no assumptions to be made about the shape of $\rho(r)$ or $\rho_{DM}(r)$, but it is computationally intensive for large N_{bin} .

The choice of N_{bin} is a compromise between accuracy in reproducing $\rho(r)$ and computational resources. Large values

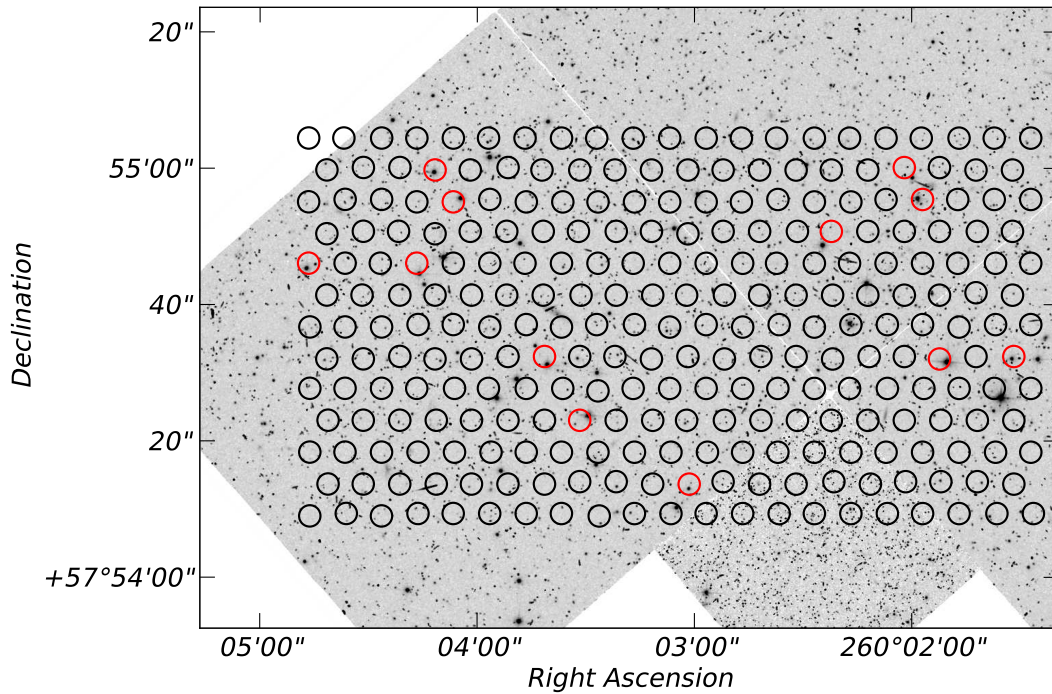


FIG. 1.— VIRUS-W IFU overlaid on top of an *HST* image from Ségall et al. (2007). Red circles highlight fibers containing stars used in the determination of the central LOSVD. Note the *HST* PSF is significantly smaller than the typical $2''$ seeing at McDonald Observatory.

of N_{bin} can make parameter space impossibly large, while small values can be overly restrictive on $\rho(r)$. We have experimented with $N_{\text{bin}} = 5, 7$, and 10 . The added freedom with $N_{\text{bin}} = 7$ or 10 was not found to be worth the increase to the dimensionality of parameter space. We have also tried connecting the ρ_i with splines, but found the additional freedom produced unrealistic density profiles. Concern over the smoothness of $\rho(r)$ may be mitigated by the fact that $\rho(r)$ only matters to our models in that it determines the potential. As the potential is the integral of $\rho(r)$, this introduces additional smoothness.

We extrapolate the density at the outermost point as a power law with slope α_∞ . The only parameters in the model are the ρ_i themselves and the extrapolation slope α_∞ . The models also have the flexibility to add a central black hole of mass M_\bullet to the galaxy for future studies.

2.2. Separating dark from stellar mass

Once the best-fitting $\rho(r)$ is found, the task remains still to recover the underlying dark matter density profile. This involves finding some other constraint on the stellar mass-to-light ratio. We can often determine M_*/L from simple stellar population (SSP) models. The required input for SSP models varies greatly, and different methods are appropriate depending on the galaxy modeled. For example, if spectra are available, stellar population synthesis models or Lick indices can be used. Lacking spectra, one can use the relations between broad-band colors and M_*/L (Bell & de Jong 2001). In nearby dSph galaxies where individual stars are resolved, color-magnitude diagrams can be constructed to fit for age and metallicity with isochrones. We can also evaluate the radial variation of M_*/L as well without much additional effort. Spectral or photometric data need only be spatially binned with the same procedure repeated at each bin. Once M_*/L is

calculated, stellar density is simply the product of the (possibly radially varying) $M_*/L \times \nu(r)$.

3. APPLICATION TO DRACO

We apply our new non-parametric Schwarzschild modeling technique to study the nearby Draco dSph. Draco is a satellite galaxy of the Milky Way orbiting at a distance from the sun of only 71 kpc (Odenkirchen et al. 2001). At this distance individual stars are resolved even with ground-based observatories. Consequently, the data we use are radial velocities of individual stars. Radial velocities are available for 158 stars in Draco, and we present radial velocities from new observations of an additional 12 stars near the center of Draco.

We choose Draco because it is the most dark matter-dominated of the “classical” (pre-SDSS) dSphs. We can therefore differentiate between dark and luminous mass more easily since the baryons contribute less to the total density profile than they do in larger galaxies. Consequently, we can absorb larger uncertainties in M_*/L . The primary science goal of this work, and a future study of all dSphs, is to determine the functional form of the dark matter profile in dSphs and compare results to theoretical predictions by CDM.

3.1. Data

3.1.1. Kinematics

We use a combination of published radial velocities and new observations for kinematics in Draco. Data exist at larger radii for 158 stars (Kleyna et al. 2002), but we wish to explore the central region of Draco in order to have the best constraint on the inner slope of $\rho_{DM}(r)$.

To accomplish this, we observe the center of Draco with the VIRUS-W integral field unit (IFU) spectrograph (Fabricius et al. 2008) on the 2.7m Harlan J. Smith telescope

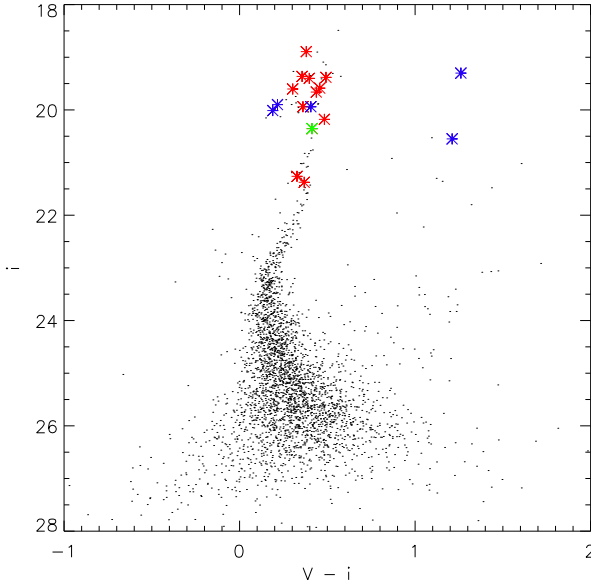


FIG. 2.— Color-magnitude diagram of stars near the center of Draco. Colored asterisks are stars we observe, coded according to their offset from Draco’s systemic velocity V_{sys} . Red stars have $|V - V_{\text{sys}}| < 30 \text{ km s}^{-1}$, blue stars have $|V - V_{\text{sys}}| > 50 \text{ km s}^{-1}$, and the green star has a radial velocity between 30 and 50 km s^{-1} of V_{sys} .

at McDonald Observatory. This instrument allows for a high density of stars to be observed simultaneously, but with the drawback that fibers are not positionable. There are 267 fibers that cover the $105'' \times 55''$ field of view with a $1/3$ fill factor. We observed the spectral region covering 4900\AA to 5500\AA with a resolving power $R \sim 9000$.

The observations took place over the first half of 5 nights from 2011 August 1-5 in excellent conditions. Seeing was typically $2''$ or better, which is smaller than the $3.2''$ diameter fibers. The standard battery of bias, Hg-Ne arc lamp, and twilight calibration frames were taken at the start of each night. We use an early implementation of the Cure data reduction software. Cure is being developed as the pipeline for the Hobby-Eberly Dark Energy Experiment (HETDEX) (Hill et al. 2006). We briefly describe steps taken to reduce the VIRUS-W data. A detailed description of Cure is beyond the scope of this paper.

We perform standard CCD processing steps, using the fitstools package (described in Gössl & Riffeser 2002), to create master bias, twilight flat, and arc lamp images for each night. We use twilight flats in combination with arc lamp images to determine the distortion solution—a two-dimensional map which translates the (x, y) position of a pixel on the CCD to a fiber number and wavelength.

Our science frames consist of 15-minute integrations of a single pointing of the central part of the galaxy. Prior to observing, we determined the optimal position of the IFU by examining *Hubble Space Telescope* (HST) photometry of the central region (Ségall et al. 2007). With accurate fiber and star positions, we determined a pointing that maximizes the number of bright stars on fibers (see Figure 1). There are 57 science frames with this pointing, totaling roughly 14 hours of integration.

We apply each night’s distortion solution to the science frames yielding rectified, wavelength-calibrated frames. We then collapse and median-combine these science frames along

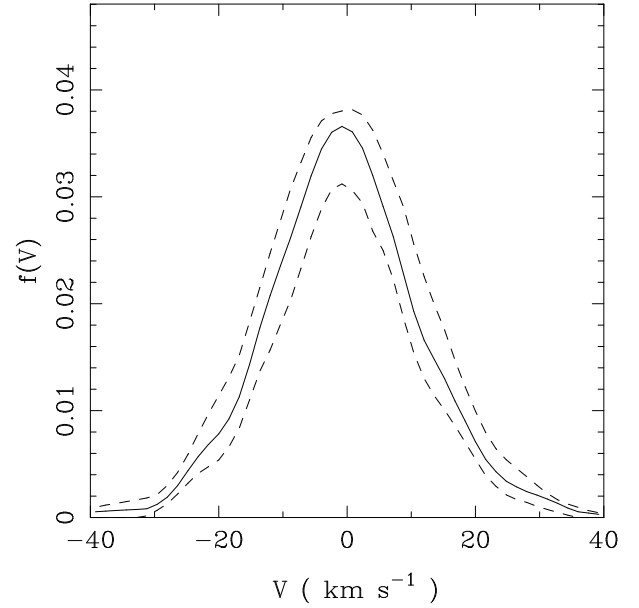


FIG. 3.— LOSVD generated from the discrete velocities of 29 stars

with the twilight flat frames. Each night’s stacked science frame is divided by the appropriate master flat for that night.

Since the majority of the 267 fibers in the IFU are on empty sky, we are able to calculate an accurate sky model directly from each science frame. We compute this sky model for each fiber on the chip using a moving-window average of 20 nearby fibers. We subtract the sky model from each frame, and the resulting sky-subtracted frames for each night are median-combined.

We extract 1-D spectra from 17 fibers containing stars. Star 2 in our sample is used as a velocity standard since it is the brightest member star with known radial velocity from Armandroff et al. (1995). We cross-correlate the other 16 spectra to star 2 using the IRAF task FXCORR. By cross-correlating to the spectrum of a star with known heliocentric radial velocity in Draco, we automatically remove the contribution from Earth-Sun motion. We perform the cross-correlation analysis on the combined image, and in doing so introduce a small bias due to the change in the heliocentric velocity correction over the course of the observing run. However, the magnitude of this change is only 0.1 km s^{-1} , much smaller than our uncertainties.

We list the heliocentric radial velocities and Tonry-Davis R_{TD} values determined for the 12 stars we report as members in Table 1. The Tonry-Davis value indicates the relative strength of the primary peak in the cross-correlation function to the average (Tonry & Davis 1979). The right ascension and declination given for each star in Table 1 indicate the position of the center of the VIRUS-W fiber containing that star.

To determine membership for the 17 stars, we use the photometry of Ségall et al. (2007) to produce a color-magnitude diagram (CMD). Figure 2 presents the resulting CMD, where the colored symbols indicate observed stars. We also group the stars according to their offset from Draco’s systemic velocity, which we assume is $V_{\text{sys}} = -293 \text{ km s}^{-1}$ (Armandroff et al. 1995). Stars with radial velocity offsets greater than 50 km s^{-1} are classified as non-members, while stars with offsets less than 30 km s^{-1} are categorized as members. The one star with radial velocity $V - V_{\text{sys}} =$

TABLE 1
RADIAL VELOCITIES OBTAINED WITH VIRUS-W

Star	RA	Dec	V_{helio} km s ⁻¹	ΔV_{helio} km s ⁻¹	R_{TD}
1	17 ^h 20 ^m 14 ^s .76	+57°54'32''40	-288.1	2.57	4.76
2	17 ^h 20 ^m 07 ^s .49	+57°54'32''04	-299.1 ¹	1.89 ¹	...
3	17 ^h 20 ^m 06 ^s .12	+57°54'32''40	-293.1	3.99	8.75
4	17 ^h 20 ^m 14 ^s .11	+57°54'23''04	-310.9	3.35	6.01
5	17 ^h 20 ^m 12 ^s .10	+57°54'13''68	-270.6	3.37	7.47
6	17 ^h 20 ^m 16 ^s .78	+57°54'59''76	-276.2	1.91	12.98
7	17 ^h 20 ^m 08 ^s .14	+57°55'00''12	-258.4	3.89	7.87
8	17 ^h 20 ^m 16 ^s .44	+57°54'55''08	-293.2	6.05	8.01
9	17 ^h 20 ^m 07 ^s .80	+57°54'55''44	-307.6	4.51	10.92
10	17 ^h 20 ^m 09 ^s .48	+57°54'50''76	-277.7	3.61	8.02
11	17 ^h 20 ^m 19 ^s .10	+57°54'46''08	-292.2	3.23	10.94
12	17 ^h 20 ^m 17 ^s .11	+57°54'46''08	-277.8	2.17	14.39

NOTE. — Heliocentric radial velocities for the 12 member stars observed with VIRUS-W at the center of Draco

¹ From Armandroff et al. (1995)

32.6 ± 3.9 km s⁻¹ (green symbol in Figure 2) is classified as a possible member. Possible and non-members are discarded from further analysis, leaving 12 member stars. Note that blind sigma-clipping retains these same 12 stars as members.

We have individual radial velocities for stars at positions around the galaxy, but our models want the distribution of radial velocities at each position—the LOSVDs. We group the individual velocities into spatial bins and determine the LOSVD at each bin via an adaptive kernel density estimator (Silverman 1986; Gebhardt et al. 1996). In velocity space, this procedure replaces each of the N discrete observations with a kernel of width h and height proportional to $N^{-1}h^{-1}$. We use the Epanechnikov kernel (an inverted parabola) and sum the contribution from each discrete velocity to obtain a non-parametric representation of the LOSVD. The 1σ uncertainties on the LOSVDs are calculated through bootstrap resamplings of the data (i.e. sampling with replacement from the N velocity measurements, see Gebhardt et al. 1996; Javelle & Gebhardt 2012). In Figure 3 we show an example LOSVD.

We combine the new VIRUS-W data with 158 additional radial velocities from the literature (Kleyna et al. 2002). We divide these 170 radial velocities into 8 radial bins of roughly 20 stars each. LOSVDs are calculated for each of these bins, yielding kinematics coverage over the radial range 25''–1500'' (8 pc–500 pc). We fit Gauss-Hermite moments to the 8 LOSVDs and plot the kinematics in Figure 4. This is only done for comparison purposes as the models fit directly to the LOSVDs. We compare the velocity dispersion as determined from the Gauss-Hermite fit with the standard deviation of the individual velocities (using the biweight scale; see Beers et al. 1990) in order to determine the best value for the smoothing width h .

The issue of foreground contamination frequently comes up in the study of dSphs using individual radial velocities. There is always the possibility that some fraction of the observed stars are members of the Milky Way. These stars would be velocity outliers and therefore artificially increase the measured velocity dispersion or, in our case, the width of the LOSVD. Fortunately, the foreground Milky Way stars are well-separated in velocity space from the Kleyna et al. (2002) sample. Contaminants are also unlikely to have colors and magnitudes that place them on the red giant branch of Draco's

color-magnitude diagram. Łokas et al. (2005) use these two constraints to estimate that there are of order 1-2 Milky Way contaminants in the entire Kleyna et al. (2002) data set.

3.1.2. Photometry

Our models are required to not only match the observed LOSVDs but also the three-dimensional luminosity density profile $\nu(r)$. The first step in obtaining $\nu(r)$ is to measure the two-dimensional surface brightness profile. We use the photometry of Ségal et al. (2007) who publish a number density profile of stars in Draco. This profile covers the radial range from 15''–2400''. We extrapolate the profile as a power law out to $R = 6000''$ by fitting a constant slope to the profile in logarithmic space. To convert the number density profile to an effective surface brightness profile, we apply an arbitrary zeropoint shift in log space until the luminosity obtained by integrating the surface brightness profile is consistent with the observed luminosity (Mateo 1998). We plot this surface brightness profile in Figure 5.

We deproject the surface brightness profile according to the procedure detailed in Gebhardt et al. (1996). We assume surfaces of constant luminosity density ν are coaxial spheroids and perform an Abel inversion. For Draco we adopt an ellipticity of $e = 0.3$ (Odenkirchen et al. 2001). We assume an inclination of $i = 90^\circ$ for simplicity. Inclination is typically one of the more difficult quantities to constrain (Thomas et al. 2007b). In addition to simplifying our models, assuming $i = 90^\circ$ provides the advantage that the deprojection is unique. For a detailed discussion of how uncertainties in viewing angle and geometry propagate through our models see Thomas et al. (2007a).

The resulting luminosity profile we calculate has an average logarithmic slope $(d \ln \nu / d \ln r) = -0.4$ inside 50 pc. In Figure 5 we plot $\nu(r)$ and also illustrate the positions of our kinematics data.

3.2. Models

Our non-parametric models of Draco use $N_{\text{bin}}=5$ radial bins spaced equally in $\log r$ from 15'' to 2000''. We initialize our search for the minimum with a brute force method, constructing a coarse grid in $N_{\text{bin}}+1$ dimensions from which we calculate all possible permutations of the N_{bin} parameters and the extrapolated slope α_∞ . Additionally, we require the density

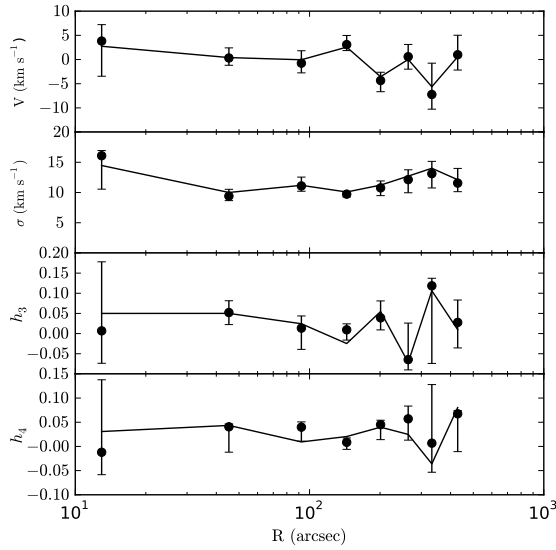


FIG. 4.— Gauss-Hermite moments fit to the 8 LOSVDs generated from 170 radial velocities. The solid line is the result of our best-fit model.

profile of each model to be monotonically decreasing or constant. This is a natural constraint, and it significantly lowers the number of models needed to sample parameter space.

Once the models defining the coarse grid are evaluated, we employ an iterative sampling scheme to focus in on and define the minimum in better detail. This method takes all the models with χ^2 within χ_{lim}^2 of the minimum χ_{min}^2 as starting points. For each starting point, a fractional step of size δ_i is taken above and below the initial value, one at a time, for all the density bins. If there is no change to χ_{min}^2 , then δ_i is decreased. This procedure is repeated until δ_i is less than a specified threshold. Additional models are also run as needed to fill in regions of parameter space that appear interesting.

We do not attempt to fit for α_∞ as we clearly do not have kinematics in that radial range to constrain the mass. Instead, we treat α_∞ as a nuisance parameter and marginalize over it in our analysis. We restrict the value of the extrapolated slope to $\alpha_\infty \in \{-2, -3, -4\}$ and every $\rho(r)$ we sample has been run with each of these values. These slopes are representative of the isothermal, NFW, and Hernquist (1990) density profiles.

Since Draco orbits within the dark matter halo of the Milky Way, it is probable that it has been tidally stripped at large radii. To account for this, the density is truncated at the tidal radius r_t defined by

$$r_t \sim \left(\frac{m}{3M}\right)^{1/3} D. \quad (4)$$

For reasonable values of the Milky Way's mass M , Draco's mass m , and the Galactocentric radius of Draco's orbit D (assumed circular), Equation 4 gives an approximate tidal radius $r_t \approx 3$ kpc. We therefore truncate $\rho(r)$ at this radius. We also assume the dark halo in Draco has the same flattening as the stars and therefore leave q_{DM} fixed at 0.7. In the future we plan to investigate models with varying q_{DM} , however that is not the focus of this paper.

4. RESULTS

The χ^2 curves for all the ρ_i are plotted in Figure 6. Each dot represents a single model, and the red curve is a smoothed fit

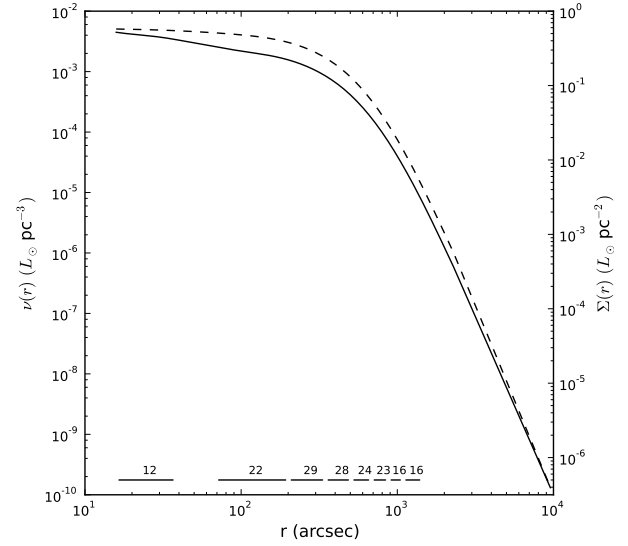


FIG. 5.— Surface brightness profile $\Sigma(r)$ (dashed) and deprojected luminosity density profile $\nu(r)$ (solid) used in our models. Horizontal lines near the x-axis indicate the radial position of our kinematics bins. Numbers refer to the number of radial velocities used per bin. Note the central location of the new VIRUS-W data (innermost bin) in comparison to existing data.

to the minimum. We obtain the red curve through a smoothing process that is similar to a boxcar average, except that we take the biweight of the 7 lowest χ^2 values within the boxcar. This method is therefore less sensitive to outliers than a traditional boxcar average. When determining a smoothed fit to the minimum, one is tempted to use only the points with the lowest χ^2 . However, numerical noise causes models to scatter to both higher and lower χ^2 in some bins. This is difficult to see by eye because scatter to higher values of χ^2 causes the models to blend in with the black points in Figure 6 while scatter to lower χ^2 makes models appear to stand out. The sliding biweight robustly picks out the center of this distribution.

The red curve plots $\chi^2(\rho_i)$ for each radial bin, and therefore gives an indication of the model-preferred density at radius r_i . We estimate the 1σ uncertainties on each of the ρ_i by determining the portion of each parameter's χ^2 curve, marginalized over all other parameters, that lies within $\Delta\chi^2 = 1$ of the overall minimum. Figure 6 shows this limit as a horizontal line whose intersection with the red curve indicates the 1σ range of the density at bin i . In all further analysis we identify the midpoint of this range as the best-fitting value and report uncertainties as symmetric about this value.

In two cases, bins 4 and 5, there are secondary minima that extend almost to the $\Delta\chi^2 = 1$ line but not quite. It is likely that with perfect coverage of parameter space the area between these minima would be filled in. However, available computational resources limit the extend to which we can sample parameter space. In order to be more conservative in our analysis we fit a quadratic in $\log \rho$ to these minima, centered roughly on the midpoint between them (blue curves in Figure 6).

The best-fitting model has unreduced $\chi_{\text{min}}^2 = 9.1$, and the number of observables our models fit to is $\nu = N_{\text{LOSVD}} \times N_{\text{vel}} = 8 \times 15 = 120$. If we were to naively calculate a reduced χ^2 , we would estimate $\chi_\nu^2 = 0.08$. This low value of χ_ν^2 results from

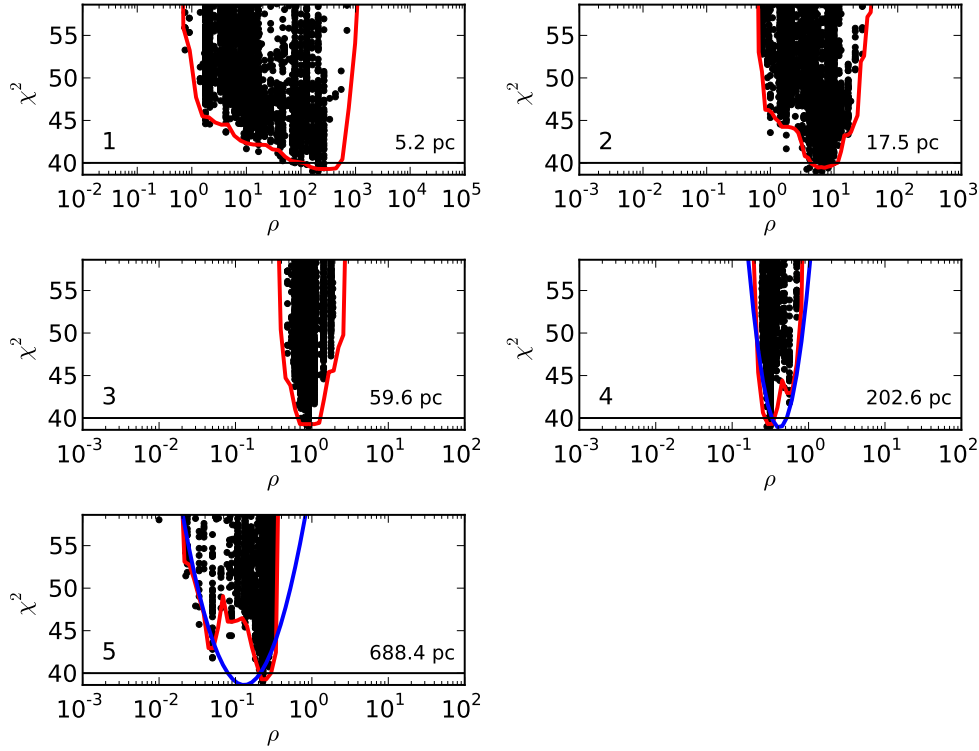


FIG. 6.— χ^2 curves for all of the ρ_i parameters. Each black dot represents a single model (combination of $\rho_1, \rho_2, \dots, \rho_5$) and the red curve is a smoothed fit to the minimum. The red curve in any panel therefore is the χ^2 curve marginalized over the other density points. The unit of density is $M_\odot \text{ pc}^{-3}$. In panels 4 and 5, the blue curve is a parabola in $\log \rho$ that we use to interpolate between two local minima.

an overestimation of the number of independent degrees of freedom ν . The adaptive kernel density estimator we use to compute the LOSVDs introduces correlations among neighboring velocity bins, therefore reducing the number of truly independent degrees of freedom.

To account for this, we consider the Gauss-Hermite parameterizations of our best-fitting model (solid line in Figure 4) and input LOSVDs (points with error bars in Figure 4). This model has $\chi_{\nu_{GH}}^2 = 0.33$ where ν_{GH} is 4 Gauss-Hermite parameters \times 8 LOSVDs = 32. This $\chi_{\nu_{GH}}^2$ is still less than 1, however it is more consistent with previous studies (Gebhardt et al. 2003) and may be due to correlations among the Gauss-Hermite parameters (e.g. Houghton et al. 2006). We use $\chi_{\nu_{GH}}^2$ to calculate the appropriate scaling to apply to our models which use the LOSVDs in determining χ^2 . We scale all un-reduced χ^2 values by the factor $\chi_{\nu_{GH}}^2 / \chi_\nu^2 = 4.3$

4.1. Obtaining M_*/L

We have so far identified the best-fitting total density profile. In order to study the dark matter profile we must subtract the stellar density profile $\rho_*(r)$. This involves finding an independent constraint on the stellar mass-to-light ratio M_*/L . Using stars within the central $5'$ of Draco, we construct a $g' - i'$ color-magnitude diagram (CMD) from the photometry of Ségal et al. (2007). We fit isochrones (Marigo et al. 2008) to the CMD, corrected for Galactic extinction (Schlegel et al. 1998), so that we may determine the age and metallicity of the stellar population.

Figure 7 shows the CMD with our best isochrone fit. The

red giant branch is well-defined, and we obtain a sensible fit with age $t_{\text{age}} = 12.7$ Gyr and metallicity $[\text{Fe}/\text{H}] = -1.4$. Using the SSP models from Maraston (2005) we are able to convert t_{age} and $[\text{Fe}/\text{H}]$ to a V-band stellar mass-to-light ratio $M_*/L_V = 2.9 \pm 0.6$. Uncertainties in M_*/L_V represent the spread in SSP predictions when different initial mass functions are assumed in the models.

4.2. The Dark Matter Profile

With M_*/L_V determined from stellar population models, we can subtract $\rho_*(r)$ from the best-fitting total density profile obtained during the modeling procedure. We plot the resulting dark matter profile in Figure 8. The red band is the 68% confidence band for each density point, marginalized over the others, and the gray band shows the 68% confidence band of all the parameters jointly (at $\Delta\chi^2 = 7.04$).

From Figure 8, it seems plausible that $\rho_{DM}(r)$ can be fit by a power law of the form $\log \rho_{DM} = \alpha \log r + \beta$ with the exception of perhaps the innermost data point. The slope of this fit α can be directly compared to both theoretical predictions and observations of similar dSphs. The innermost point, however, is puzzling. Its value indicates a large central density and a departure from the power-law nature of the outer profile. Further puzzling is that its point-wise uncertainty (plotted as a red error bar) indicates strong constraint despite the fact that we have no kinematic data in this region of the galaxy. We speculate that, in the absence of such data, models are able to arbitrarily increase the central density. Since the volume of this inner bin is small, the total amount of mass added is negligible. With no kinematics in this region, models can easily

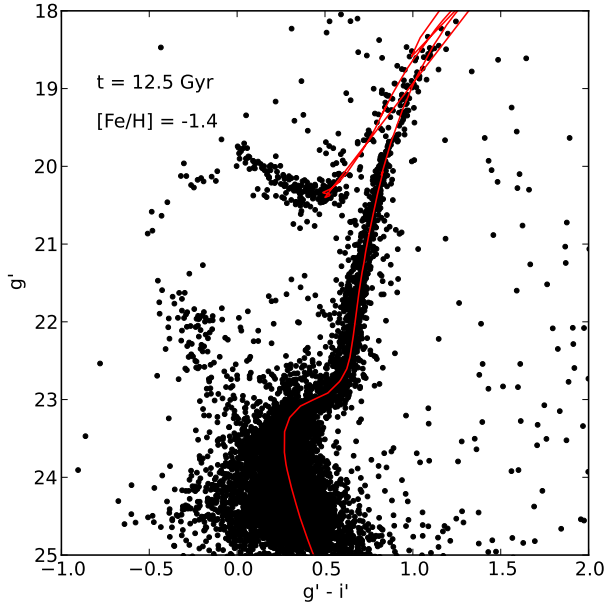


FIG. 7.— Color-magnitude diagram of stars within the central 5' of Draco. From left to right, we plot isochrones of $(t_{\text{age}} \times 10^9 \text{ yr}, [\text{Fe}/\text{H}]) = (11.5, -1.6)$, $(12.5, -1.4)$, and $(13.5, -1.3)$. The solid red line is the $(12.5, -1.4)$ isochrone we use when determining M_*/L_V .

absorb this mass without affecting χ^2 . We therefore exclude the innermost point in all further analysis.

The resulting power law fit to the outer four points is shown in blue in Figure 8. We characterize the uncertainty in this fit by constructing 1,000 Monte Carlo realizations with noise added to the density profile. We draw each point point i randomly from a Gaussian distribution with mean $\log \rho_i$ and dispersion given by the width of the 1σ confidence band at point i in Figure 8. We repeat the fit for each realization, and determine the 1σ uncertainties on α from the 68% span of this distribution. This procedure yields $\alpha = -1.0 \pm 0.2$. None of the 1,000 realizations has a slope $\alpha > -0.45$ strongly indicating that the galaxy is not cored for $r \gtrsim 20$ pc.

4.3. Orbit Structure

Once we have determined the best-fitting model, we can calculate the internal (unprojected) moments of the distribution function at each of the bins in our meridional grid. Of interest is the anisotropy in the velocity dispersion tensor, which we quantify with the ratio σ_r/σ_t —the ratio of radial to tangential anisotropy in the galaxy. We define the tangential anisotropy σ_t as

$$\sigma_t \equiv \sqrt{\frac{1}{2}(\sigma_\theta^2 + \sigma_\phi^2 + v_\phi^2)} \quad (5)$$

in spherical polar coordinates where v_ϕ is the rotational velocity. Streaming motions in the r and θ directions are assumed to be zero. We plot σ_r/σ_t in Figure 9. Since the LOSVDs we use in Draco contain contributions from stars at all angles θ , we average σ_r and all quantities in Equation (5) when calculating σ_r/σ_t . Consequently, we lose the ability to evaluate anisotropy as a function of θ . This can be avoided if better kinematics coverage is available, either through more stars with radial velocities in dSphs or two-dimensional integral-

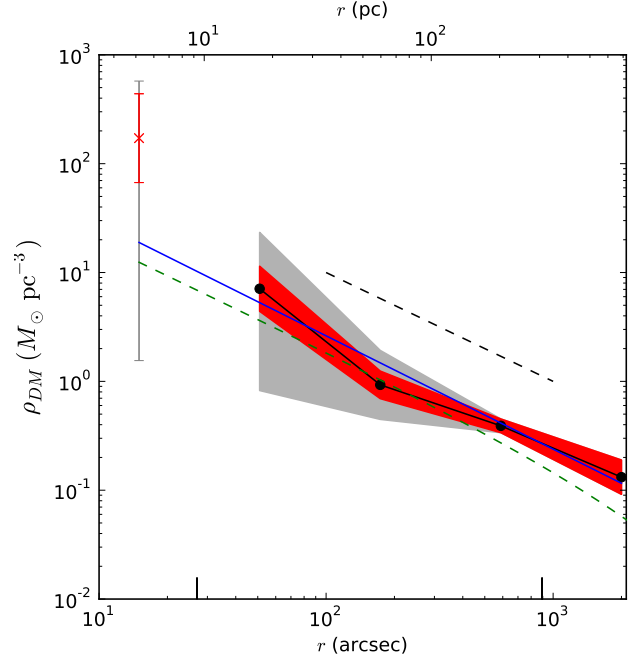


FIG. 8.— Best-fitting dark matter density profile in Draco. The red shaded region represents the point-wise 68% confidence band for $\rho_{DM}(r)$ ($\Delta\chi^2 = 1$), with the solid black line derived from forcing symmetric logarithmic errors. The gray shaded region is the 68% confidence band on $\rho_{DM}(r)$ considering all parameters jointly ($\Delta\chi^2 = 7.04$). We plot the innermost point (excluded from all further analysis) an error bar with the same color scheme. The solid blue line is the best power law fit to the profile, and the dashed line shows an r^{-1} NFW-like profile. We plot the best-fitting NFW halo from a small grid of parametric models as the dashed green line. Vertical lines along the x-axis indicate the radial range of our kinematic data.

field spectroscopy in more distant galaxies. Fortunately most other large dSphs in the Local Group have many more radial velocities publicly available.

We plot σ_r/σ_t in Figure 9 over the radial range that our LOSVDs sample. We determine the uncertainties in σ_r/σ_t by the maximum/minimum values of σ_r/σ_t for models within $\Delta\chi^2 = 7.04$ of χ^2_{\min} (1σ for $N_{\text{bin}}+1$ degrees of freedom). We find evidence for radial anisotropy at all radii, consistent with the “tidal stirring” theory describing the origin of the Milky Way dSphs (Łokas et al. 2010; Kazantzidis et al. 2011). Uncertainties are large on σ_r/σ_t , likely due to the small number of radial velocities available as kinematic constraint. To constrain the anisotropy better, more radial velocities are needed.

5. DISCUSSION

5.1. Improvement over Parametric Methods

Since we eventually fit our non-parametric dark matter profile with a power law, one can ask why we do not initially use a power law-parameterized profile. This would seem advantageous, especially given the large parameter space required by non-parametric methods. This reasoning, however, relies on the assumption that we know the profile is a power law a priori. The point of this study is to relax this assumption and see what type of profile comes out of the modeling, rather than impose unjustified interpretation on the problem. It happens that Draco hosts a nearly power law density profile, but by not assuming this a priori we allow more general models to be explored. As a rough check that our models have converged to a global minimum, we run a small grid of parametric models with an NFW dark matter density profile. The best-fitting of

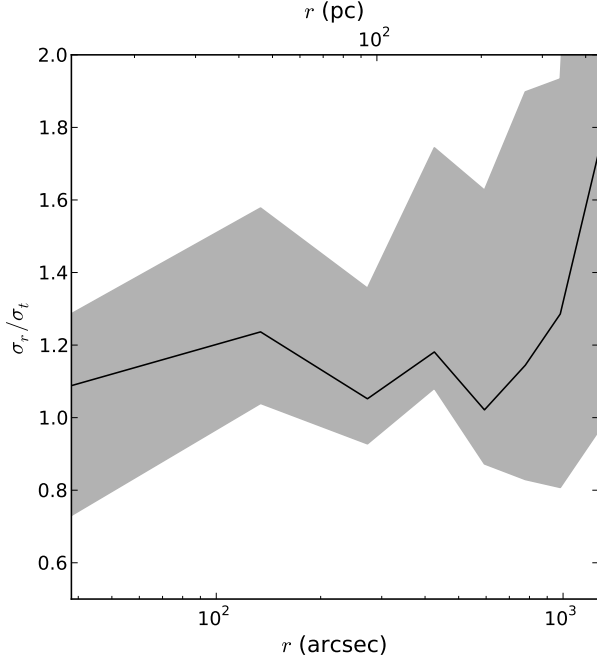


FIG. 9.— Ratio of the radial to tangential components of the velocity dispersion. Values of σ_r/σ_t different from unity indicate anisotropy. The black line is our best-fitting model.

these models is plotted in green in Figure 8.

5.2. Interpreting the Dark Matter Profile

It is important to note that we only constrain the dark matter density profile over little more than a decade in radius from 20–700 pc. One could easily imagine our power law fit changing from $\alpha = -1$ to a core ($\alpha = 0$) inside of $r \sim 20$ pc. Likewise, the slope may also change at larger radii than $r \sim 700$ pc without our knowledge. The NFW density profile has an outer slope $\alpha = -3$ for $r \gg r_s$, but our profile does not change slope within our model grid. This could indicate that $r_s \gg 700$ pc, but without knowledge of the outer slope we cannot say with certainty that the profile is NFW-like.

Recent cosmological N -body simulations have been found to produce density profiles shallower than the traditional $\alpha = -1$ cusps (Stadel et al. 2009; Navarro et al. 2010). Many authors suggest that dark matter profiles are best parameterized by the Einasto profile (Navarro et al. 2004; Merritt et al. 2005; Gao et al. 2008; Navarro et al. 2010) where the slope varies with radius according to a power law $\alpha(r) \propto r^n$. These profiles can have shallower cusps than NFW, but do not have constant slopes over a large range in radius. Our non-parametric density profile is well-fit by a single power law from $20 \lesssim r \lesssim 700$ pc, but, again, this is a fairly narrow range in radius. Our models cannot rule out an Einasto-like change in slope outside this radial range. More kinematics are needed to characterize the density profile at large and small radii.

When calculating the potential, we allow the outer slope of $\rho(r)$ to vary between $2 \leq \alpha_\infty \leq 4$ for $r > 700$ pc, but, unsurprisingly, we are unable to constrain α_∞ . Tidal effects may also alter the shape of $\rho_{DM}(r)$ since Draco is orbiting within the dark matter halo of the Milky Way. The tidal radius calculated from Equation (4) is sufficiently large that tides are unlikely to affect the stellar component, but $\rho_{DM}(r)$ at large radii could be affected. If this is the case, $\rho_{DM}(r)$ would decline more steeply than expected and the total mass enclosed

would be smaller than what we calculate.

The cuspy $\alpha = -1$ dark matter profile we find in Draco stands in contrast to many other observational studies of dSphs that find $\alpha = 0$ cores (Gilmore et al. 2007; Walker & Peñarrubia 2011; Jardel & Gebhardt 2012). The effects of baryons are still not well-understood, and could potentially drive α to different values on a galaxy-by-galaxy basis. These effects are the sum of at least two competing processes. Adiabatic compression (Blumenthal et al. 1986) draws in dark matter boosting the central ρ_{DM} and driving α to more negative values. On the other hand, feedback from star formation and supernovae can cause strong outflows (Navarro et al. 1996a; Binney et al. 2001) which can in turn remove dark matter from the centers of galaxies, reshaping cuspy profiles into $\alpha = 0$ cores.

In a recent paper, Governato et al. (2012) use high resolution cosmological N -body simulations with a fully hydrodynamical treatment of baryons to test these two competing effects in low-mass dwarf galaxies. They find that the cusps of the dark matter halo is directly related to the amount of star formation activity in the galaxy. This is expressed as a correlation between α and stellar mass M_* . Their interpretation is that galaxies that form more stars (larger M_*) have more supernovae and a greater potential to turn a cuspy dark matter profile into a core. Using their least-squares fit to the M_* - α correlation, they predict $\alpha \approx -1.3$ (at 500 pc) for Draco’s stellar mass. This is in approximate agreement with our measured value of $\alpha = -1$.

Perhaps owing to the lack of stellar velocities available in Draco compared to other dSphs, there are not many studies investigating its dark matter profile through dynamical models. A rough comparison can be made with Łokas et al. (2005) who fit profiles with an inner slope of $\alpha = -1$ and an outer exponential cutoff at large radii. They find a total mass-to-light ratio that varies with radius between $100 \gtrsim M_{\text{tot}}/L_V \gtrsim 1000$ in the inner ~ 700 pc. These values are comparable to the total mass-to-light ratio we calculate in the inner ~ 300 pc. However, unlike Łokas et al. (2005) we do not impose an exponential cutoff in $\rho_{DM}(r)$ at large radii. Our calculated M_{tot}/L_V therefore rises sharply at large radii where the stellar luminosity profile is decreasing much faster than $\rho_{DM}(r)$.

Importantly, $M_{\text{tot}}/L_V \gg M_*/L_V = 2.9 \pm 0.6$ (the stellar mass-to-light ratio we derive from SSP models) at all radii. This means we can confidently state that Draco is dark matter-dominated at all radii, allowing us to easily absorb errors in M_*/L_V from SSP models. In other words, when determining $\rho_{DM}(r)$ from Equation (1) the uncertainty in $\rho(r)$ dominates the uncertainty in stellar density since the product $M_*/L \times \nu(r)$ is much smaller than $\rho(r)$. This is one of the reasons we choose to test this non-parametric technique on Draco first. In the future we plan to extend this analysis to the remaining Local Group dSphs, which are also thought to be dark matter-dominated everywhere.

5.3. Draco’s Mass

We plot the enclosed mass profile of our models in Figure 10. The shaded region is the 1σ confidence band derived from the extreme values of $M(r)$ for all models within $\Delta\chi^2 = 5.84$ of the minimum (1σ for $N_{\text{bin}} = 5$ free parameters, marginalizing over α_∞). The vertical ticks on the x-axis represent the radial extent of our kinematics coverage. From this plot it is apparent that, despite its low luminosity and stellar mass, Draco lives in a dark matter halo that is surprisingly massive.

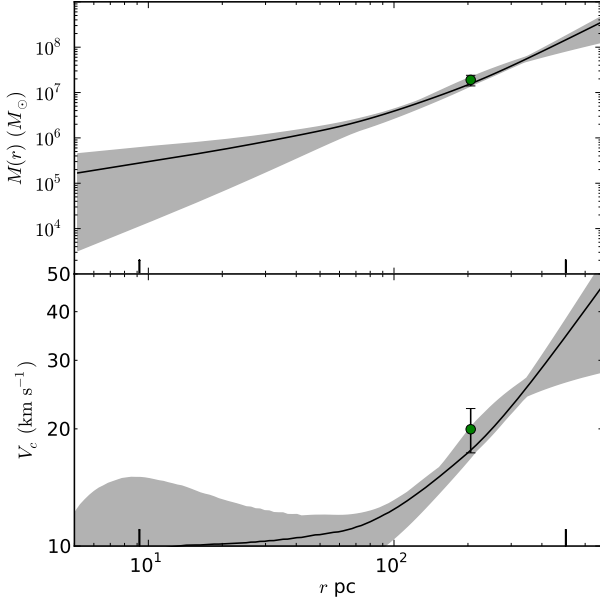


FIG. 10.— (Top): Enclosed Mass profile of our best-fitting model (black line) and 1σ confidence region. The green point is the Wolf et al. (2010) mass estimator. (Bottom): Circular speed profile and 1σ confidence region. Colors are the same as above. Vertical tick marks on the x-axis represent the range of our kinematics coverage.

An interesting comparison can be made with the brightest dSph Fornax, roughly two orders of magnitude higher in luminosity. If we compare the mass enclosed within a common physical radius of 300 pc, we find that for Draco $M_{300} \equiv M(r=300 \text{ pc}) = 3.8^{+0.84}_{-0.29} \times 10^7 M_\odot$, and Jardel & Gebhardt (2012) measure $M_{300} = 3.5^{+0.77}_{-0.11} \times 10^6 M_\odot$ for Fornax. Of course, Fornax is much more extended than Draco so it is sensible to also compare the mass enclosed within the deprojected half-light radius of each galaxy’s stellar component. For Draco we measure $M_{1/2} \equiv M(r=r_e) = 1.6^{+0.6}_{-0.2} \times 10^7 M_\odot$, and in Fornax Jardel & Gebhardt (2012) measure $M_{1/2} = 5.8^{+1.0}_{-0.2} \times 10^7 M_\odot$. We would prefer to compare the total mass of each galaxy, but there are no kinematic tracers far enough out in the halo that the density profile declines sharply enough to keep mass finite for any dSph. Consequently, we cannot constrain the total mass observationally and we must rely on comparisons to simulations (Section 5.4).

We also use our dynamical models to compare our measurement of $M_{1/2}$ with the convenient mass estimator proposed by Wolf et al. (2010) (see Walker et al. 2009 and Cappellari et al. 2006 for similar formulae). This formula relates $M_{1/2}$ to the directly observable luminosity-weighted line-of-sight velocity dispersion $\langle \sigma_{LOS}^2 \rangle$ and projected half-light radius R_e . The Wolf et al. (2010) mass estimator is written as:

$$M_{1/2} \approx 4G^{-1}R_e \langle \sigma_{LOS}^2 \rangle \quad (6)$$

and Wolf et al. (2010) give a theoretical argument for why the effect of anisotropy is minimized near r_e for a variety of stellar systems in spherical symmetry.

For a more fair comparison of Equation (6) to our models we calculate $M_{1/2}$ from our data set, not the value listed in Wolf et al. (2010). We use $\langle \sigma_{LOS}^2 \rangle = 11.3 \pm 1.6 \text{ km s}^{-1}$, calculated directly from our data in Figure 4, as well as

$R_e = 158.1 \text{ pc}$ and $r_e = 205.2 \text{ pc}$ which we derive from the photometry in Figure 5. This calculation yields an estimated $M_{1/2} = (1.9 \pm 0.5) \times 10^7 M_\odot$, in excellent agreement with the mass calculated from our models. We plot the estimated $M_{1/2}$ as the green point in Figure 10.

5.4. Comparing Draco to CDM Simulations

We can also gain insight into the properties of Draco’s dark matter halo by examining the circular speed profile $V_c(r)$ plotted in the lower panel of Figure 10. The green point plotted is $V_{1/2} = \sqrt{GM_{1/2}/r_{1/2}} = 20.0 \pm 2.6 \text{ km s}^{-1}$ using our value of the Wolf et al. (2010) mass estimator. In a recent paper, Boylan-Kolchin et al. (2012) match the observed $V_{1/2}$ of Local Group dSphs to subhalos around a Milky Way-like halo in the Aquarius Simulation (Springel et al. 2008) to derive constraints on each dSph’s maximum circular speed V_{max} —a quantity directly related to the total halo mass. Boylan-Kolchin et al. (2012) find that this estimate of V_{max} is usually $20 - 30 \text{ km s}^{-1}$ smaller than the V_{max} they obtain through abundance matching. These results lead them to conclude that the Local Group dSphs are dynamically inconsistent with the types of halos they are predicted to inhabit from abundance matching.

We are in a position to investigate this claim directly in Draco. We do not need to match our $V_{1/2}$ to simulations in order to gain knowledge of $V_c(r)$; we calculate the latter directly, and not just at the half-light radius. Interestingly, much of our circular speed profile lies above the $V_{\text{max}} = 20.5^{+4.8}_{-3.9}$ predicted by Boylan-Kolchin et al. (2012). At $r = 500 \text{ pc}$, the radius where we run out of kinematic tracers and can therefore no longer robustly constrain the mass, we find $V_c = 34.6^{+3.5}_{-8.2} \text{ km s}^{-1}$. We can take the lower bound of V_c here as lower limit on $V_{\text{max}} \geq 26.4$. The scaling relations between total mass and V_{max} for subhalos (Springel et al. 2008) imply a lower limit on Draco’s total mass of $M \geq 1.0 \times 10^9 M_\odot$.

Ours is not the first study to suggest that Draco lives in a halo with such a large mass. Peñarrubia et al. (2008) demonstrate that a family of NFW halos with varying V_{max} and r_{max} are consistent with the stellar kinematics of any King model embedded in an NFW halo. They break this degeneracy by invoking the correlation between V_{max} and r_{max} found in CDM simulations (e.g. Bullock et al. 2001). Their study suggests that Draco is the most massive of the Milky Way dSphs with $V_{\text{max}} \approx 35 \text{ km s}^{-1}$.

The comparison between Draco and Fornax is interesting as the two galaxies are separated by almost two orders of magnitude in luminosity but may have similar masses. Since Draco’s inner halo is nicely fit by the NFW density profile (Figure 8), we can rely on simulations to extrapolate a total mass $M \geq 1.0 \times 10^9 M_\odot$. However, multiple independent studies using different methods suggest that Fornax does not live in an NFW halo (Goerdt et al. 2006; Walker & Peñarrubia 2011; Jardel & Gebhardt 2012), and we therefore should not use the NFW formalism to predict a total mass from its V_{max} . Still, the similarity in the galaxies’ values of $M_{1/2}$ and M_{300} suggests that the simplest abundance matching models, which require a one-to-one mapping between luminosity and total mass, may not appropriately describe the dSphs. If Draco and Fornax do indeed have similar masses, despite vastly different baryonic properties, then there must be substantial stochasticity in the galaxy formation process at the dSph mass scale. Even without comparing to Fornax, it is clear that Draco’s baryonic properties do not map in the expected way to its halo

mass.

K.G. acknowledges support from NSF-0908639. We thank Dave Doss and the staff at McDonald Observatory for observing support. We gratefully acknowledge fruitful discussions

with Louis Strigari, Remco van den Bosch, Glenn van de Ven, and Andrea Maccio. This work would not be possible without the state-of-the-art computing facilities available at the Texas Advanced Computing Center (TACC).

REFERENCES

- Adams, J. J., Gebhardt, K., Blanc, G. A., et al. 2012, *ApJ*, 745, 92
- Armandroff, T. E., Olszewski, E. W., & Pryor, C. 1995, *AJ*, 110, 2131
- Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *AJ*, 100, 32
- Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212
- Binney, J., Gerhard, O., & Silk, J. 2001, *MNRAS*, 321, 471
- Blais-Ouellette, S., Amram, P., & Carignan, C. 2001, *AJ*, 121, 1952
- Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, *ApJ*, 301, 27
- Borriello, A., & Salucci, P. 2001, *MNRAS*, 323, 285
- Boylan-Kolchin, M., Bullock, J. S., & Kaplinghat, M. 2012, *MNRAS*, 2657
- Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, *MNRAS*, 321, 559
- Burkert, A. 1995, *ApJL*, 447, L25+
- Cappellari, M., Bacon, R., Bureau, M., et al. 2006, *MNRAS*, 366, 1126
- Cretton, N., de Zeeuw, P. T., van der Marel, R. P., & Rix, H. 1999, *ApJS*, 124, 383
- de Blok, W. J. G., McGaugh, S. S., Bosma, A., & Rubin, V. C. 2001, *ApJL*, 552, L23
- Evans, N. W., An, J., & Walker, M. G. 2009, *MNRAS*, 393, L50
- Fabricius, M. H., Barnes, S., Bender, R., et al. 2008, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7014, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*
- Gao, L., Navarro, J. F., Cole, S., et al. 2008, *MNRAS*, 387, 536
- Gebhardt, K., Richstone, D., Ajhar, E. A., et al. 1996, *AJ*, 112, 105
- Gebhardt, K., Richstone, D., Kormendy, J., et al. 2000, *AJ*, 119, 1157
- Gebhardt, K., Richstone, D., Tremaine, S., et al. 2003, *ApJ*, 583, 92
- Gilmore, G., Wilkinson, M. I., Wyse, R. F. G., et al. 2007, *ApJ*, 663, 948
- Goerdt, T., Moore, B., Read, J. I., Stadel, J., & Zemp, M. 2006, *MNRAS*, 368, 1073
- Gössl, C. A., & Riffeser, A. 2002, *A&A*, 381, 1095
- Governato, F., Zolotov, A., Pontzen, A., et al. 2012, *ArXiv e-prints*
- Hernquist, L. 1990, *ApJ*, 356, 359
- Hill, G. J., MacQueen, P. J., Tufts, J. R., et al. 2006, *Proc. SPIE*, 6269
- Houghton, R. C. W., Magorrian, J., Sarzi, M., et al. 2006, *MNRAS*, 367, 2
- Jardel, J. R., & Gebhardt, K. 2012, *ApJ*, 746, 89
- Kazantzidis, S., Łokas, E. L., Callegari, S., Mayer, L., & Moustakas, L. A. 2011, *ApJ*, 726, 98
- Kleyna, J., Wilkinson, M. I., Evans, N. W., Gilmore, G., & Frayn, C. 2002, *MNRAS*, 330, 792
- Koch, A., Kleyna, J. T., Wilkinson, M. I., et al. 2007, *AJ*, 134, 566
- Łokas, E. L., Kazantzidis, S., Klimentowski, J., Mayer, L., & Callegari, S. 2010, *ApJ*, 708, 1032
- Łokas, E. L., Mamon, G. A., & Prada, F. 2005, *MNRAS*, 363, 918
- Maraston, C. 2005, *MNRAS*, 362, 799
- Marigo, P., Girardi, L., Bressan, A., et al. 2008, *A&A*, 482, 883
- Mateo, M. L. 1998, *ARA&A*, 36, 435
- Merritt, D., Navarro, J. F., Ludlow, A., & Jenkins, A. 2005, *ApJL*, 624, L85
- Navarro, J. F., Eke, V. R., & Frenk, C. S. 1996a, *MNRAS*, 283, L72
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996b, *ApJ*, 462, 563
- Navarro, J. F., Hayashi, E., Power, C., et al. 2004, *MNRAS*, 349, 1039
- Navarro, J. F., Ludlow, A., Springel, V., et al. 2010, *MNRAS*, 402, 21
- Odenkirchen, M., Grebel, E. K., Harbeck, D., et al. 2001, *AJ*, 122, 2538
- Peñarrubia, J., McConnachie, A. W., & Navarro, J. F. 2008, *ApJ*, 672, 904
- Persic, M., Salucci, P., & Stel, F. 1996, *MNRAS*, 281, 27
- Richstone, D. O., & Tremaine, S. 1988, *ApJ*, 327, 82
- Rix, H., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, *ApJ*, 488, 702
- Salucci, P., & Burkert, A. 2000, *ApJL*, 537, L9
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- Schwarzschild, M. 1979, *ApJ*, 232, 236
- Ségall, M., Ibata, R. A., Irwin, M. J., Martin, N. F., & Chapman, S. 2007, *MNRAS*, 375, 831
- Silverman, B. W. 1986, *Density estimation for statistics and data analysis*, ed. Silverman, B. W.
- Simon, J. D., Bolatto, A. D., Leroy, A., Blitz, L., & Gates, E. L. 2005, *ApJ*, 621, 757
- Siopis, C., Gebhardt, K., Lauer, T. R., et al. 2009, *ApJ*, 693, 946
- Springel, V., Wang, J., Vogelsberger, M., et al. 2008, *MNRAS*, 391, 1685
- Stadel, J., Potter, D., Moore, B., et al. 2009, *MNRAS*, 398, L21
- Strigari, L. E., Bullock, J. S., Kaplinghat, M., et al. 2008, *Nature*, 454, 1096
- Thomas, J., Jesseit, R., Naab, T., et al. 2007a, *MNRAS*, 381, 1672
- Thomas, J., Saglia, R. P., Bender, R., et al. 2005, *MNRAS*, 360, 1355
- . 2007b, *MNRAS*, 382, 657
- . 2004, *MNRAS*, 353, 391
- Tonry, J., & Davis, M. 1979, *AJ*, 84, 1511
- Valluri, M., Merritt, D., & Emsellem, E. 2004, *ApJ*, 602, 66
- van den Bosch, R. C. E., van de Ven, G., Verolme, E. K., Cappellari, M., & de Zeeuw, P. T. 2008, *MNRAS*, 385, 647
- van der Marel, R. P., Cretton, N., de Zeeuw, P. T., & Rix, H. 1998, *ApJ*, 493, 613
- Walker, M. G., Mateo, M., Olszewski, E. W., et al. 2009, *ApJ*, 704, 1274
- Walker, M. G., & Peñarrubia, J. 2011, *ApJ*, 742, 20
- Wolf, J., Martinez, G. D., Bullock, J. S., et al. 2010, *MNRAS*, 406, 1220